

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/526,479  
Source: PGT/10  
Date Processed by STIC: 3/13/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/526,479

DATE: 03/13/2005

TIME: 12:09:38

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03132005\J526479.raw

3 <110> APPLICANT: Duncan McGREGOR  
 4 Richard ODEGRIP  
 5 Kevin FITZGERALD  
 6 Rosemarie HEDERER  
 7 Bill ELDRIDGE  
 8 Chris ULLMAN  
 9 Philip KUHLMAN  
 10 Davod COOMBER  
 12 <120> TITLE OF INVENTION: IN VITRO PEPTIDE EXPRESSION LIBRARY  
 14 <130> FILE REFERENCE: 117-533 / N.86234C SER  
 C--> 16 <140> CURRENT APPLICATION NUMBER: US/10/526,479  
 17 <141> CURRENT FILING DATE: 2005-03-04  
 19 <150> PRIOR APPLICATION NUMBER: PCT/GB03/03860  
 20 <151> PRIOR FILING DATE: 2003-09-05  
 22 <150> PRIOR APPLICATION NUMBER: GB 0304657.0  
 23 <151> PRIOR FILING DATE: 2003-02-28  
 25 <150> PRIOR APPLICATION NUMBER: GB 0304521.8  
 26 <151> PRIOR FILING DATE: 2003-02-27  
 28 <150> PRIOR APPLICATION NUMBER: GB 0220759.5  
 29 <151> PRIOR FILING DATE: 2002-09-06  
 31 <160> NUMBER OF SEQ ID NOS: 48  
 33 <170> SOFTWARE: MS Word  
 36 <210> SEQ ID NO: 1  
 37 <211> LENGTH: 23  
 38 <212> TYPE: DNA  
 39 <213> ORGANISM: Artificial sequence  
 41 <220> FEATURE:  
 42 <223> OTHER INFORMATION: Primer  
 44 <400> SEQUENCE: 1  
 45 actgatcttc accaaacgta tta 23  
 48 <210> SEQ ID NO: 2  
 49 <211> LENGTH: 22  
 50 <212> TYPE: DNA  
 51 <213> ORGANISM: Artificial sequence  
 53 <220> FEATURE:  
 54 <223> OTHER INFORMATION: Primer  
 56 <400> SEQUENCE: 2  
 57 tgcatactctg tctgtccaca gg 22  
 60 <210> SEQ ID NO: 3  
 61 <211> LENGTH: 48  
 62 <212> TYPE: DNA  
 63 <213> ORGANISM: Artificial sequence  
 65 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 03/13/2005

PATENT APPLICATION: US/10/526,479

TIME: 12:09:38

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03132005\J526479.raw

```

66 <223> OTHER INFORMATION: Primer
68 <400> SEQUENCE: 3
69 gagcttcaac aggggagggg gaggaggatc aactgatctt caccaaac 48
72 <210> SEQ ID NO: 4
73 <211> LENGTH: 50
74 <212> TYPE: DNA
75 <213> ORGANISM: Artificial sequence
77 <220> FEATURE:
78 <223> OTHER INFORMATION: Primer
80 <400> SEQUENCE: 4
81 ctaggactgg attcaacggg gggaggagga tcaactgac ttcaccaaac 50
84 <210> SEQ ID NO: 5
85 <211> LENGTH: 49
86 <212> TYPE: DNA
87 <213> ORGANISM: Artificial sequence
89 <220> FEATURE:
90 <223> OTHER INFORMATION: Primer
92 <400> SEQUENCE: 5
93 cagaagagga tctgaatggg ggaggagggt ccactgtggc tgcaccatc 49
96 <210> SEQ ID NO: 6
97 <211> LENGTH: 22
98 <212> TYPE: DNA
99 <213> ORGANISM: Artificial sequence
101 <220> FEATURE:
102 <223> OTHER INFORMATION: Primer
104 <400> SEQUENCE: 6
105 tcccctgttg aagctctttg tg 22
108 <210> SEQ ID NO: 7
109 <211> LENGTH: 40
110 <212> TYPE: DNA
111 <213> ORGANISM: Artificial sequence
113 <220> FEATURE:
114 <223> OTHER INFORMATION: Primer
116 <400> SEQUENCE: 7
117 cagaagagga tctgaatggg ggaggagggt ccggaaaacc 40
120 <210> SEQ ID NO: 8
121 <211> LENGTH: 27
122 <212> TYPE: DNA
123 <213> ORGANISM: Artificial sequence
125 <220> FEATURE:
126 <223> OTHER INFORMATION: Primer
128 <400> SEQUENCE: 8
129 gctacgttga atccagtcct aggagag 27
132 <210> SEQ ID NO: 9
133 <211> LENGTH: 22
134 <212> TYPE: DNA
135 <213> ORGANISM: Artificial sequence
137 <220> FEATURE:
138 <223> OTHER INFORMATION: Primer

```

## RAW SEQUENCE LISTING

DATE: 03/13/2005

PATENT APPLICATION: US/10/526,479

TIME: 12:09:38

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03132005\J526479.raw

```

140 <400> SEQUENCE: 9
141 catattgtcg ttagaacgcg gc 22
144 <210> SEQ ID NO: 10
145 <211> LENGTH: 58
146 <212> TYPE: DNA
147 <213> ORGANISM: Artificial sequence
149 <220> FEATURE:
150 <223> OTHER INFORMATION: Primer
152 <400> SEQUENCE: 10
153 attcagatcc tcttctgaga tgagtttttg ttcctcgagc atggtagatc ctgtttcc 58
156 <210> SEQ ID NO: 11
157 <211> LENGTH: 42
158 <212> TYPE: DNA
159 <213> ORGANISM: Artificial sequence
161 <220> FEATURE:
162 <223> OTHER INFORMATION: Primer
164 <400> SEQUENCE: 11
165 cgatacctag cgttcggatc catattgtcg ttagaacgcg gc 42
168 <210> SEQ ID NO: 12
169 <211> LENGTH: 1788
170 <212> TYPE: DNA
171 <213> ORGANISM: Artificial sequence
173 <220> FEATURE:
174 <223> OTHER INFORMATION: DNA construct
176 <400> SEQUENCE: 12
177 catattgtcg ttagaacgcg gctacaatta atacataacc ttatgtatca tacacatacg 60
178 atttaggtga cactatagaa tacaagctta ctcccatcc cctgttgac aattaatcat 120
179 ggctcgtata atgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acaggatcta 180
180 ccatgctcga ggaacaaaaa ctcatctcag aagaggatct gaatggggga ggagggtcca 240
181 ctgtggctgc accatctgtc ttcattctcc cgccatctga tgagcagttg aaatctggaa 300
182 ctgcctctgt tgtgtgcctg ctgaataact tctatccag agaggccaaa gtacagtggg 360
183 aggtggataa cgccctccaa tcgggtaact ccaggagag tgtcacagag caggacagca 420
184 aggacagcac ctacagcctc agcaacaccc tgacgctgag caaagcagac tacgagaaac 480
185 acaaagtcta cgctgcgaa gtcacccatc agggcctgag ctgcccgtc acaaagagct 540
186 tcaacagggg agggggagga ggatcaactg atcttcacca aacgtattac cgccaggtaa 600
187 agaaccgaa tccggtgttc actccccgtg aaggtgccgg aacgccgaag ttccgcgaaa 660
188 aaccgatgga aaaggcgggtg ggcctcacct cccgttttga ttccgccatt catgtggcgc 720
189 atgcccgttc ccgtggtctg cgtcggcgca tgccaccggt gctgcgtcga cgggctattg 780
190 atgcgctgct gcaggggctg tgtttccact atgaccgct ggccaaccgc gtccagtgtt 840
191 ccatcaccac actggccatt gagtgcggac tggcgacaga gtccggtgca ggaaaactct 900
192 ccatcaccgc tgccaccggt gccctgacgt tctgtcaga gctgggactg attacctacc 960
193 agacggaata tgaccgctt atcgggtgct acattccgac cgacatcacg ttcacactgg 1020
194 ctctgtttgc tgcccttgat gtgtctgagg atgcagtggc agctgcgcgc cgcagtcgtg 1080
195 ttgaatggga aaacaaacag cgcaaaaagc aggggctgga taccctgggt atggatgagc 1140
196 tgatagcgaa agcctggcgt tttgtgcgtg agcgtttccg cagttaccag acagagcttc 1200
197 agtcccgtgg aataaaacgt gcccgctgcgc gtcgtgatgc gaacagagaa cgtcaggata 1260
198 tcgtcaccct agtgaaacgg cagctgacgc gtgaaatctc ggaaggacgc ttcactgcta 1320
199 atggtgaggc ggtaaaacgc gaagtggagc gtcgtgtgaa ggagcgcgtg attctgtcac 1380
200 gtaaccgcaa ttacagccgg ctggccacag cttctccctg aaagtgatct cctcagaata 1440

```

## RAW SEQUENCE LISTING

DATE: 03/13/2005

PATENT APPLICATION: US/10/526,479

TIME: 12:09:38

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03132005\J526479.raw

```

201 atccggcctg cgccggaggc atccgcacgc ctgaagcccg ccggtgcaca aaaaaacagc 1500
202 gtcgcatgca aaaaacaatc tcatcatcca ccttctggag catccgattc cccctgtttt 1560
203 taatacaaaa tacgcctcag cgacggggaa ttttgcttat ccacatttaa ctgcaagggg 1620
204 cttccccata aggttacaac cgttcatgtc ataaagcgcc agccgccagt cttacagggg 1680
205 gcaatgtatc ttttaaacac ctgtttatat ctcttttaaa ctacttaatt acattcattt 1740
206 aaaaagaaaa cctattcact gcctgtcctg tggacagaca gatatgca 1788

```

209 &lt;210&gt; SEQ ID NO: 13

210 &lt;211&gt; LENGTH: 1518

211 &lt;212&gt; TYPE: DNA

212 &lt;213&gt; ORGANISM: Artificial sequence

214 &lt;220&gt; FEATURE:

215 &lt;223&gt; OTHER INFORMATION: DNA construct

217 &lt;400&gt; SEQUENCE: 13

```

218 catattgtcg ttagaacgcg gctacaatta atacataacc ttatgtatca tacacatacg 60
219 atttaggtga cactatagaa tacaagctta ctccccatcc cctgtttgac aattaatcat 120
220 ggctcgtata atgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acaggatcta 180
221 ccattgctga ggaacaaaaa ctcatctcag aagaggatct gaatggggga ggagggtccg 240
222 gaaaacctat cccaaacctt ctctaggac tggattcaac ggggggagga ggatcaactg 300
223 atcttcacca aacgtattac cgccaggtaa agaaccgaa tccggtgttc actccccgtg 360
224 aaggtgccgg aacgccgaag ttccgcgaaa aaccgatgga aaaggcgggt ggccctacct 420
225 cccgttttga tttgcctatt catgtggcgc atgccggtc ccggtgtctg cgtcggcgca 480
226 tggcaccggg gctgcgtcga cgggctattg atgcgctgct gcaggggctg tgtttccact 540
227 atgaccgcgt ggccaaccgc gtccagtgtt ccatcaccac actggccatt gagtgcggac 600
228 tggcgacaga gtccggtgca ggaaaactct ccatcaccgc tggcaccggt gccctgacgt 660
229 tcctgtcaga gctgggactg attacctacc agacggaata tgaccgctt atcgggtgct 720
230 acattccgac cgacatcacg ttcacactgg ctctgtttgc tgcctttgat gtgtctgagg 780
231 atgcagtggc agctgcgcgc cgcagtcgtg ttgaatggga aaacaaacag cgcaaaaagc 840
232 aggggctgga taccctgggt atggatgagc tgatagcgaa agcctggcgt tttgtgcgtg 900
233 agcgtttccg cagttaccag acagagcttc agtcccgtgg aataaaacgt gccctgcgc 960
234 gtcgtgatgc gaacagagaa cgtcaggata tcgtcaccct agtgaaacgg cagctgacgc 1020
235 gtgaaatctc ggaaggacgc ttcactgcta atggtgaggc ggtaaaacgc gaagtggagc 1080
236 gtcgtgtgaa ggagcgcatt attctgtcac gtaaccgcaa ttacagccgg ctggccacag 1140
237 cttctccctg aaagtgatct cctcagaata atccggcctg cgccggaggc atccgcacgc 1200
238 ctgaagcccg ccggtgcaca aaaaaacagc gtcgcatgca aaaaacaatc tcatcatcca 1260
239 cttctctggag catccgattc cccctgtttt taatacaaaa tacgcctcag cgacggggaa 1320
240 ttttgcttat ccacatttaa ctgcaagggg cttccccata aggttacaac cgttcatgtc 1380
241 ataaagcgcc agccgccagt cttacagggg gcaatgtatc ttttaaacac ctgtttatat 1440
242 ctcttttaaa ctacttaatt acattcattt aaaaagaaaa cctattcact gcctgtcctg 1500
243 tggacagaca gatatgca 1518

```

246 &lt;210&gt; SEQ ID NO: 14

247 &lt;211&gt; LENGTH: 38

248 &lt;212&gt; TYPE: DNA

249 &lt;213&gt; ORGANISM: Artificial sequence

251 &lt;220&gt; FEATURE:

252 &lt;223&gt; OTHER INFORMATION: Estrogen Receptor Target Recognition Sequence

254 &lt;400&gt; SEQUENCE: 14

```

255 tcaggtcaga gtgacctgag ctaaaataac acattcag 38

```

258 &lt;210&gt; SEQ ID NO: 15

259 &lt;211&gt; LENGTH: 828

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/526,479

DATE: 03/13/2005

TIME: 12:09:38

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03132005\J526479.raw

```

260 <212> TYPE: DNA
261 <213> ORGANISM: Artificial sequence
263 <220> FEATURE:
264 <223> OTHER INFORMATION: repA sequence
266 <220> FEATURE:
267 <221> NAME/KEY: CDS
268 <222> LOCATION: (1)..(828)
269 <223> OTHER INFORMATION:
W--> 271 <400> 15
272 atg gta aag aac ccg aat ccg gtg ttc act ccc cgt gaa ggt gcc gga      48
273 Met Val Lys Asn Pro Asn Pro Val Phe Thr Pro Arg Glu Gly Ala Gly
274 1          5          10          15
276 acg ccg aag ttc cgc gaa aaa ccg atg gaa aag gcg gtg ggc ctc acc      96
277 Thr Pro Lys Phe Arg Glu Lys Pro Met Glu Lys Ala Val Gly Leu Thr
278          20          25          30
280 tcc cgt ttt gat ttc gcc att cat gtg gcg cat gcc cgt tcc cgt ggt      144
281 Ser Arg Phe Asp Phe Ala Ile His Val Ala His Ala Arg Ser Arg Gly
282          35          40          45
284 ctg cgt cgg cgc atg cca ccg gtg ctg cgt cga cgg gct att gat gcg      192
285 Leu Arg Arg Arg Met Pro Pro Val Leu Arg Arg Ala Ile Asp Ala
286          50          55          60
288 ctg ctg cag ggg ctg tgt ttc cac tat gac ccg ctg gcc aac cgc gtc      240
289 Leu Leu Gln Gly Leu Cys Phe His Tyr Asp Pro Leu Ala Asn Arg Val
290 65          70          75          80
292 cag tgt tcc atc acc aca ctg gcc att gag tgc gga ctg gcg aca gag      288
293 Gln Cys Ser Ile Thr Thr Leu Ala Ile Glu Cys Gly Leu Ala Thr Glu
294          85          90          95
296 tcc ggt gca gga aaa ctc tcc atc acc cgt gcc acc cgg gcc ctg acg      336
297 Ser Gly Ala Gly Lys Leu Ser Ile Thr Arg Ala Thr Arg Ala Leu Thr
298          100          105          110
300 ttc ctg tca gag ctg gga ctg att acc tac cag acg gaa tat gac ccg      384
301 Phe Leu Ser Glu Leu Gly Leu Ile Thr Tyr Gln Thr Glu Tyr Asp Pro
302          115          120          125
304 ctt atc ggg tgc tac att ccg acc gac atc acg ttc aca ctg gct ctg      432
305 Leu Ile Gly Cys Tyr Ile Pro Thr Asp Ile Thr Phe Thr Leu Ala Leu
306          130          135          140
308 ttt gct gcc ctt gat gtg tct gag gat gca gtg gca gct gcg cgc cgc      480
309 Phe Ala Ala Leu Asp Val Ser Glu Asp Ala Val Ala Ala Ala Arg Arg
310 145          150          155          160
312 agt cgt gtt gaa tgg gaa aac aaa cag cgc aaa aag cag ggg ctg gat      528
313 Ser Arg Val Glu Trp Glu Asn Lys Gln Arg Lys Lys Gln Gly Leu Asp
314          165          170          175
316 acc ctg ggt atg gat gag ctg ata gcg aaa gcc tgg cgt ttt gtg cgt      576
317 Thr Leu Gly Met Asp Glu Leu Ile Ala Lys Ala Trp Arg Phe Val Arg
318          180          185          190
320 gag cgt ttc cgc agt tac cag aca gag ctt cag tcc cgt gga ata aaa      624
321 Glu Arg Phe Arg Ser Tyr Gln Thr Glu Leu Gln Ser Arg Gly Ile Lys
322          195          200          205
324 cgt gcc cgt gcg cgt cgt gat gcg aac aga gaa cgt cag gat atc gtc      672

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/526,479

DATE: 03/13/2005  
TIME: 12:09:39

Input Set : A:\Sequence Listing.txt  
Output Set: N:\CRF4\03132005\J526479.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; N Pos. 29,30,32,33,35,36,38,39,41,42,44,45,47,48,50,51,53,54,56,57

Seq#:22; N Pos. 59,60,62,63

Seq#:25; N Pos. 695,696,698,699,701,702,704,705,707,708,710,711,713,714,716

Seq#:25; N Pos. 717,719,720,722,723,725,726,728,729

Seq#:28; N Pos. 37,38,40,41,43,44,46,47,49,50,52,53,55,56,58,59,61,62,64,65

Seq#:28; N Pos. 67,68,70,71,73,74,76,77,79,80,82,83,85,86,88,89

## VERIFICATION SUMMARY

DATE: 03/13/2005

PATENT APPLICATION: US/10/526,479

TIME: 12:09:39

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\03132005\J526479.raw

L:16 M:270 C: Current Application Number differs, Wrong Format  
L:271 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:15,Line#:269  
L:545 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0  
M:341 Repeated in SeqNo=22  
L:692 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:660  
M:341 Repeated in SeqNo=25  
L:846 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:0  
M:341 Repeated in SeqNo=28